Author Responses for "Lifecycle of Updrafts and Mass Flux in Isolated Deep Convection over the Amazon Rainforest: Insights from Cell Tracking" (egusphere-2023-2410)

The authors would like to thank the referees for their insightful comments and suggestions to improve the quality of the manuscript. This document has referee comments in blue, author responses in black, and "text added to the revised manuscript in red within quotations".

REFEREE 1 (RC1):

The authors present novel, insightful and pertinent research into the relation between convective intensity, lifecycle, season and convective mass flux using a combination of RWP and geostationary satellite observations with a cloud-tracking approach. This research is highly relevant as an improved understanding of how convective-scale processes impact convective mass flux, and how these processes change over the convective lifecycle, is vital to understanding their connected to larger scale changes such as anvil area and lifetime. In addition, the study is performed using data from the Amazon region, where deep convective dynamics are important but factors affecting convective-scale processes are less well understood.

In general I find the paper to be insightful and provide useful information, and the results are generally well described. Clear conclusions are reached, and these are linked to prior research. However, I have several concerns regarding how the lifecycle stages of the tracked DCCs are defined. These lifecycle stages are crucial to the further analysis of the paper. In addition, while the results are generally described clearly in the text, the figures are less clear. I would like to see improvements in how they are presented, and the inclusion of additional figures to show how features are detected using the radar data. My recommendation is that the article be reconsidered after major revisions.

Major comments:

Section 2.4: It is very unclear exactly how the lifetime stages are defined. There are several references to the stages being "determined" by Z, cell area or their tendencies, but it seems from L208-211 that the actual definition of the five-lifetime stages is quintiles of the total cell lifetime. Is this correct? The labelling of these bins is then justified by the average Z and area of features within each lifetime quintile. L213 justifies the third quintile as being the mature stage as it has the highest average Z and area, although figure 3c shows that the 4th bin has the largest average area. Have the authors considered whether the different lifecycle stages make up the same proportion of cell lifetime in all cases? In particular, are longer-lived cells expected to have the same proportions as shorter-lived cells?

We have modified the text in Section 2.4 to avoid confusion. Our decision making is now explicitly described to guide the reader through the steps of defining the lifecycle stages of convection. We start by dividing the data into 5 bins of normalized lifetime values. We then associate these bins with lifecycle stages of convection based on the bulk statistical trends in Z and cell area. This assignment is done following previously defined lifecycle stages of convection (e.g., Futyan and Del Genio, 2007) that have been used extensively in the literature. Regarding the highest cell area, on average, the 3rd and 4th lifetime bins had average values of cell area that were within 1 km^2 with statistically insignificant differences. Therefore, we have kept the current lifecycle stage definitions given the higher average Z for the 3rd lifetime bin. The 4th lifetime bin is defined as the late-mature stage as the cell area remains large. This choice is justified in later sections as the results indicate the 3rd lifetime bin represents stronger convection with higher rainfall rates and stronger updrafts compared to other stages. The proportion of shorter-lived versus longer-lived cells in each lifetime bin is discussed in response to the comment below.

The second major issue with section 2.4 is a potential sampling bias on the lifecycle bins for cell lifetime due to the similar number of bins as the average number of observations of most features. In particular, the most common observed lifetime of 36 minutes means that there are four radar scans of the tracked cell. In this case, therefore, there will always be one lifecycle bin that is never classified for cells of this length. Based on a quintile normalisation, that will be bin 3 or the “mature” stage. If we assume that longer lived cells are also more intense, then we would expect to have a bias in the Z of “mature” DCC observations because short-lived DCCs are never classified into bin 3. Is this the case, and if so, could the authors look into whether this does affect the analysis? I am not sure how to resolve this issue, other than reducing the number of bins or increasing the minimum length of cells selected for analysis.

It is true that cells with a lifetime value of 36 minutes do not contribute to the 3rd lifetime bin. However, any impact of this bias on the analysis depends on two assumptions: (1) There are substantial differences in the average lifetime of cells that contribute data to each lifetime bin and (2) any differences in convection intensity based on cell lifetime also affect the bulk statistics of the convection properties and the lifecycle determination. This is because the remainder of the analysis involving updraft properties and mass flux does not utilize the SIPAM radar data. We found that neither assumption is true for this dataset and the classification bias does not affect the analysis (detailed discussion is in the following paragraphs). Based on this finding, we have not changed the lifecycle definitions and include cells with a lifetime of 36 minutes in our analysis to preserve the sample size of the study.

There are minor differences in the average lifetime for cells that contribute data to each lifetime bin. The average lifetime of cells contributing to each lifetime bin is 92, 93, 100, 95, and 90 mins, respectively. These differences are less than 10 mins regardless of whether cells with a lifetime of 36 mins are included or excluded. If cells with lifetime values of 36 mins were to be excluded, the average lifetime of cells contributing to each bin would be 101, 104, 100, 108, and 98 mins, respectively. As highlighted below, these differences are small enough to not lead to significant changes in convection properties if shorter lived cells are excluded (discussed below).

We quantified the impact of the lifecycle classification on the bulk statistics of the average radar reflectivity (Z) as a proxy for convection intensity. This is done by comparing the average Z for each lifetime bin before and after removing cells with lifetime values of 36 mins. The average Z increases by less than 0.2 dBZ for each bin (except the 3rd bin) when cells with lifetime values of

36 mins are excluded. Since these shorter-lived cells contribute about 20 % of the sample size, we do not exclude these cases from the analysis. Their inclusion does not cause a substantial change in the statistics when the average propagation speed or cell area are considered.

We refrain from directly comparing longer-lived cells versus shorter-lived cells in this study. This is because the cell lifetimes are not evenly distributed as shown in Fig. 3a, c. The number of cells is heavily skewed toward cells with shorter lifetimes and the number of cells decreases rapidly as the cell lifetime values increase. This is expected due to the nature of the isolated convection targeted in this study and the temporal resolution of the SIPAM radar. The inverse relationship between cell lifetime and the number of cells is more pronounced for isolated convection compared to organized or widespread deep convection.

The proportion of cells with a particular range of lifetime values does not vary significantly across the five bins (except when considering cells with lifetime = 36 mins, in which case the 3rd lifetime bin does not have any contribution). We have added this table to the supplement as Table S2.

Table S2: Proportion of cells (in percentage) with different cell lifetimes in each lifetime bin.

Lifetime (L) in minutes	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
$36 \leq L \leq 60$	34	37	32	32	38
$60 < L \leq 120$	46	43	46	46	44
$120 < L \leq 240$	18	19	19	20	17
$240 < L \leq 360$	1.7	1.8	2.0	2.2	1.4

We have added the following text to mention the classification bias in the manuscript and the lack of its impact on the analysis: "The proportional contribution of cells with different cell lifetime values to the five lifecycle stages or bins is provided in Table S2. The lifecycle classification described above meant that cells with a lifetime of 36 minutes do not contribute data to the 3rd lifetime bin. This is because there are five lifetime bins while these shorter-lived cells consist of four radar scans without any normalized lifetime value corresponding to the 3rd lifetime bin. However, this does not affect our analysis or the bulk statistics of the convection properties. This is because the average lifetime of cells contributing data to each bin is within 10 mins regardless of whether these shorter-lived cells are included. The similarity in the average cell lifetime across the bins is a result of an equally proportional distribution of cells lasting for 36 mins or more into the five lifetime bins (Table S2). Consequently, the average SIPAM Z for each lifetime bin changes by less than 0.2 dBZ depending on whether the shorter-level cells were included. To preserve our sample size, we include these shorter-lived cells in the subsequent analyses. The choice of these lifecycle definitions and the data classification is justified by results presented in the following sections."

Figures: In general I found that the figures were not very clear, particularly from figure 8 onwards. Use of similar colours for the lines in figures 8 onwards makes them difficult to distinguish compared to the line colours used in figures 6 & 7. In addition, multiple different lines are plotted in the same charts without using different line styles (e.g. dashed, dotted) to distinguish them. In many plots, "lifetime bin" is used as the x axis. It would be clearer to a reader looking first at the figures if this was changed to "lifetime stage" with labels "developing", "early mature", "mature" etc. In addition, several of the figures show the same properties but for the wet or dry seasons (figures 9 & 10, 14 & 14). It would be easier to compare these cases if both wet & dry season plots were included on the same figure. In addition, I think it would help the reader to understand the features being tracked if an example of the gridded Z field with tracked features was included as an additional figure.

We have added new figures and updated the existing figures to address these suggestions:

1. We added snapshots of the gridded Z field with tracked features from June 22, 2014 during 15:00 to 15:24 Z in Figure 1.

We have added the following description of the snapshots in Section 2.2: "Figure 1 shows an example of the gridded reflectivity field near MAO and reflectivity features identified by *tobac* between 15:00 and 15:24 Z on June 22, 2014. Markers represent the position of each feature with Z > 30 dBZ and polygons surrounding these features represent their areal extent based on the segmentation process."

2. We changed the line colors starting with Figure 8 from the original manuscript in consistency with the colors used in Figures 6 and 7.
3. We used different line styles or translucency to better distinguish between different lines in the same panels.
4. We updated the figures showing data from dry and wet seasons to show these seasonal contrasts in a single figure.
5. We kept the original x-axis labels of 'lifetime bins' for figures 6 and 7 from the original manuscript since these labels highlight what is directly plotted in these figures. The text provides the reader with context to associate these bins to the corresponding lifecycle stages.

Minor comments:

L35: It may be informative to include a more recent reference showing that km-scale convective resolving models do not fully resolve this issue

We have added a reference to "Prein et al., 2021". The missing citation has been added to the list of references. We have also added a reference to "Ramos-Valle et al., 2023" where the authors simulate convective systems over the Amazon region using Weather Research and Forecasting model simulations at different grid spacings.

L71: Jeyaratnam et al., 2020 missing from references

There was a typo as we meant to cite "Jeyaratnam et al., 2021". The missing citation has been added to the list of references.

L77: Minor: The acronym for 'Manacapuru, Brazil' is listed here as T3/MAO, but MAO is used throughout the text. It would be clearer simply to list the acronym as MAO here.

We have changed this to "MAO".

L111: How much is the gridded Z affected by low-level convection and congestus at 2km?

Gridded Z at 2 km is affected by both low-level convection and congestus clouds. However, these data are removed from the analysis following the criteria outlined in Section 2.5:

1. Congestus clouds are screened out by using an 8 km minimum threshold for the Echo Top Height of 10 dBZ echoes from the RWP.
2. Cases of low-level convection that do not transition into deep convection are removed by removing cases when Z > 20 dBZ was not observed at altitudes between 6 to 8 km.

L120: For reproducibility it would be good to state which version of tobac was used. Also, if a newer version of tobac is used than that presented in the 2019 paper, it may be appropriate to also cite Sokolowsky & Freeman et al. 2023 (https://egusphere.copernicus.org/preprints/2023/egusphere-2023-1722/, in review)

We have added the following text to Section 2.2: "For this study, cell tracking was conducted using *tobac* 1.3.3 following improvements in linking features along a cell trajectory (Sokolowsky et al., 2023)."

L124: It would be good to also discuss previous radar cloud tracking, possibly noting that much of the application of cloud tracking with cloud radar data has focused on nowcasting rather than weather (e.g. Wilson et al. 1998 [https://doi.org/10.1175/1520-0477(1998)079%3C2079:NTASR%3E2.0.CO;2], Keenan et al. 2003 [https://doi.org/10.1175/BAMS-84-8-1041]), although more recently there has been increased focus on its use for studying convective cloud processes (e.g. Feng et al. 2022 [https://doi.org/10.1175/MWR-D-21-0237.1])

We have added the following text to the first paragraph in Section 2.2: "Geostationary satellites relate cloud-top properties of convective clouds to their evolution. For instance, satellite data from the infrared channel can be used to infer vertical cloud development based on a decrease in brightness temperatures and the cloud expansion rate based on the divergence at upper levels.

In comparison, radars can infer the content and characteristics of hydrometeors and retrieve information about shallow precipitating clouds. Radar-based studies can also associate the development and decay of the precipitating core with the rate of mass flow and precipitation within the core. Previous studies have used radar data to track the evolution of convective systems for decades, not only for operational nowcasting purposes (Wilson et al., 1998, Keenan et al., 2003, Wilhelm et al., 2023), but also for studying convective cloud processes (Rosenfeld, 1987; Wapler, 2017; Feng et al. 2022, Giangrande et al., 2023)."

L126: It would be informative to state why this improved makes a difference, both in terms of the tracking of DCCs and the study of convective processes

We have added the following text to highlight these points: "Given that less than 50 % of the cells sampled for more than 36 mins were also sampled for more than 60 mins (Table 1), the use of SIPAM radar data helped increase the study's sample size. Observations with better spatial resolution also improve the characterization of cloud processes as the profiles of vertical velocity and convective mass flux can vary within 5 to 10 mins (Fig. S1 and S2)."

L130: It would be clearer to state that these thresholds have been chosen by the authors, and link with the rest of the paragraph about why they have been chosen, and what storm features/intensity they correspond to

The current sentence states "*tobac* uses multiple thresholds (30, 40, 50, and 60 dBZ) for feature identification." We have modified this sentence to the following: "*tobac* allows users to select multiple thresholds for feature identification. For this study, thresholds of 30, 40, 50, and 60 dBZ were chosen."

L153: Initiation of the tracked cell, convection and precipitation are not necessarily the same. It is expected for there to be some delay between the initiation of convection and the radar reflectivity reaching the detection threshold. It would be better to be clear that the measured initiation is the first time of detection.

The following sentence was added: "However, it is important to note that the initial detection of Z > 30 dBZ by the radar may not reflect the exact timing of precipitation initiation."

L173: This median nearest neighbour distance seems very small. Is this the distance between the centre of each feature or the edges? Can neighbouring cells be clearly defined with this separation on a 1km grid? It may be helpful to include a figure of an example of the gridded Z field with detected features to show this clearly.

The nearest neighbor distance is determined as the distance between the position of two features. The position of each feature refers to the center of mass of the weighted reflectivity, as defined in Section 2.2. We have added snapshots of the gridded Z field with tracked features from June 22, 2014 during 15:00 to 15:24 Z in Figure 1.

L182: A 36-minute lifetime should correspond to 4, rather than 3, radar scans (inclusive of start and end time)

We have modified the text as follows: "Cells with lifetime less than 36 minutes correspond to three or fewer radar scans or feature designations."

L190: Text says 1,130 cells, but table 1 says 672 cells. Is there a difference in the selection criteria here?

We have modified the text as follows: "Based on these criteria, 1,130 cells were identified. The bulk statistics and trends in convection properties are examined for these cells. In certain instances, a radar scan had more than one feature satisfying every criteria. For such cases, the feature closest to MAO was selected to assign a lifecycle stage to the RWP data without ambiguity. This resulted in the selection of 2803 features (from 672 cells) with one feature representing each radar scan to assign a lifecycle stage to the RWP data from the radar scan timestep. The RWP data are further screened in Section 2.5 to avoid the inclusion of congestus clouds or incompletely sampled updraft cores above the MAO site. This screening meant that only 31% of the DCCs and 12% of the features initially selected were used for the analysis of RWP data in Section 3".

L244: Giangrande et al., 2016 mentions two different classification schemes (Steiner et al. 1995 & Giangrande et al., 2013). It would be helpful to provide a brief description of the scheme in the paper, rather than simply providing a reference.

We apologize for the confusion. There was a typo as we meant to cite "Giangrande et al., 2013". L244 refers to the RWP classification. The classifications schemes mentioned in the comment are applied to different instruments - Steiner et al. (1995) is applied to the SIPAM radar CAPPI, while Giangrande et al. (2013) is applied to the RWP data.

L273 I think it would be clearer if cells with areas > 500 km^2 were excluded in the criteria listed earlier, and figure 3 shown without their influence.

This figure is Figure 4 in the revised manuscript. We have updated the figure to include data from cells with areas < 500 km^2. In addition, we have moved the text describing the removal of these larger cells from our dataset to Section 2.3.

L304: I would like to see this result included in the main paper, rather than the supplementary materials. The difference in the lifetime and organisation of cells between the wet and dry seasons is a nice result to show from this analysis. Possibly also shown cell area distributions, as this is also mentioned as being different in the text

We appreciate the referee's suggestion. We decided to place Figure S3 in the revised supplementary materials given the similarities between Fig. S3 and Fig. 3 in the revised manuscript. This decision was also influenced by the high number of figures in the revised manuscript. We agree that these results are important and have provided a thorough quantitative assessment of this figure in the manuscript.

L325: These are interesting results. It would also be interesting to show if there is a difference in average cell lifetime for the different diurnal cycle bins

We have added the following text to this section: "The average cell lifetime decreased with time of day with values of 103, 95, and 91 mins for pre-sunrise, morning, and afternoon cells, respectively. This is likely due to the stronger precipitation in proportion with higher average Z leading to hydrometeor unloading and cloud depletion."

L497: These areas do not seem to match those shown in figure 3

These areas correspond to a subset of the cells represented in Figure 3. Figure 3 examines bulk statistics for all cells satisfying the criteria in Section 2.4. However, these cells were further screened to correspond with the RWP data selected in Section 2.5. This difference was highlighted on Line 251 of the original manuscript.

L622: It may also be worth noting the relative lifecycle of the anvil cloud compared to the core. i.e. we only expect to see the anvil start to dissipate after the core has dissipated, and so while the core is in its dissipating phase the anvil is still mature and therefore expected to continue to have cold T_b

Following the comment from referee #2 and considering the length of the revised manuscript with the addition of text and figures in response to referee comments, we have decided to remove section 3.4 from the revised manuscript.

L645: Figure 17c appears to show that T_b reached a maximum during the mature phase for cells with ETH between 10 and 12 km. Are the classification by T_b and ETH for each bin in b and c respectively just based on the measurements of those properties during that lifecycle bin? If so, that may confound the analysis shown here. For example, only the most intense cells are likely to have an ETH > 10km during the developing phase, and so the average T_b of these cells is expected to be cold. However, more, less intense, cells may reach the 10 km ETH level during the mature phase, which could make the average T_b warmer simply because the distribution of the intensity of sampled cells has changed. It would be clearer if cells were classified by the minimum T_b and maximum ETH observed over their entire lifetime, rather than during the lifecycle stage, for figure 17b and c respectively. However, while this should be possible for T_b, I'm not sure if this is possible with ETH provided by the RWP. Perhaps the SIPAM measurement could be used instead to provide ETH?

Following the comment from referee #2 and considering the length of the revised manuscript with the addition of text and figures in response to referee comments, we have decided to remove section 3.4 from the revised manuscript.

L649: Cases with T_b this low, particularly during the mature stage, seem like they could be due to a misidentification or mislocation of the GOES-13 imagery. Have the authors looked at what is shown in the GOES imagery here (i.e. are there colder pixels near to but not at the collocated location).

Following the comment from referee #2 and considering the length of the revised manuscript with the addition of text and figures in response to referee comments, we have decided to remove section 3.4 from the revised manuscript.

L651: Use of the 241 K threshold dates all the way back to Maddox, 1980 [https://www.jstor.org/stable/26221473], but it was selected on account of the -32C contour being provided by the GOES-1 cold T_b enhancement algorithm rather than a specific physical basis and has stuck ever since.

Following the comment from referee #2 and considering the length of the revised manuscript with the addition of text and figures in response to referee comments, we have decided to remove section 3.4 from the revised manuscript.

Figure 1: The cell tracks plotted in a/c are very unclear. It may be clearer to plot a "heatmap" with all of the tracks plotted in the same colour but with translucency. Including an arrow or similar to show the direction of each track would also be informative. In addition, the river looks like a cell track at first glance. Showing the location of the SIPAM would also be informative

We have updated Figure 1 (a, c) to remove the river from the map. We have marked the locations of the MAO site and the SIPAM radar in these panels. Following a comment from referee #2, we have added heatmaps of the probability density of the cell tracks in Figure 2. We have avoided adding the arrows to the tracks for clarity. We believe the wind rose diagrams in Figure 2 illustrate the bulk statistics of the propagation speed and direction for each season.

We have added the following text in Section 2.3: "In Figure 1a and 1b, we highlight the cell tracks for isolated DCCs from the wet (December to April) and dry season (June to September). In Fig. 2, we show wind rose diagrams based on the propagation direction of these cells and heatmaps from a 2D histogram of latitude and longitude pairings from the cell tracks."

Figure 7: Mention in the caption that the 18-00 LT bin is not shown due to the small number of samples as mentioned in the text.

We have added the following text to the figure caption: "The 18-00 LT bin is not shown due to the small number of samples."

Table S1: Please format these more clearly as dates

We have modified the table caption to the following to avoid confusion: "Table S1: List of ISO days from 2014 and 2015 in the "YYYYMMDD" format."

REFEREE 2 (RC2):

Review of egusphere-2023-2410: Lifecycle of Updrafts and Mass Flux in Isolated Deep Convection over the Amazon Rainforest: Insights from Cell Tracking

Summary: In this manuscript, the authors calculate a variety of statistics as a function of storm lifetime and season from a database of tracked convective cells from GoAmazon2014-15 observations. The authors present a commendable set of different methods, observations and statistics to provide additional insights about convective dynamics, how they vary as a function of storm lifecycle and season, and how their results fit in with recent results from their research group. Given the complexity of their analysis and dataset, there were many aspects of their analysis that warrant additional description and details to properly understand what their data sample represents and the subsequent results that are being presented. Similarly, many of the figures / analyses require additional descriptions to interpret and understand what is being presented. Therefore, while the work fits well within ACP's focus and could add insights for understanding convective storm dynamics, I am not able to fully assess the science results without a better understanding of the methods.

Major comments, questions, and/or concerns:

Combining MAO data with storms. If a storm is within 20 km of the MAO site at any point in its lifecycle, the RWP data from MAO are used and attributed to a storm lifecycle stage based on the tracking. The average distance between the feature center and MAO site was 8.5 km, with 70% of feature positions being with 10 km of the MAO site (L252). Whether the storm updraft goes directly over the site versus if its center is ~10 km away from the site might make a significant difference in the RWP data (i.e., Z, W) that are being associated with the storm (i.e., profiles of Z and W at the center of the storm updraft region will be different from the edges of the storm). This seems like a large source of uncertainty that is not discussed. For example, are there biases in the distance of the updraft center to what was sampled by MAO with respect to the different lifecycle stages and seasons, and if so, how do these biases impact the results in the manuscript? Perhaps, showing Figure 1 as a heat map both for all storms and for storms at the different stages and seasons would help address this concern.

In Figure 2, we have added heatmaps of the probability density of the lat-lon tracks for the dry and wet season, respectively. We found that about 39% and 33% of the tracks for the dry and wet season, respectively, are within 0.1-degree latitude and longitude of the MAO site. About 80% and 77% of the tracks for the dry and wet season, respectively, are within 0.2-degree latitude and longitude of the MAO site. This suggests there were minor differences across seasons in terms of the distance of the cell tracks from the MAO site. Furthermore, the intent of cell tracking was not to identify the precise location of the updraft center. Instead, cell tracking was conducted to determine the age of storms when they were sampled at the MAO site by the RWP. The selection of cells closest to the MAO site is done to confidently attribute a single lifecycle stage to each updraft core sampled by the RWP. The properties of these convective cores are evaluated independent of the SIPAM radar data.

While it is challenging to confidently assess if the RWP sampled the center of the storm versus the edge of the storm, we believe the strict selection criteria applied in Section 2.5 sufficiently ensure the RWP sampled a significant portion of the storm with adequate updraft or downdraft core sampling. We developed these selection criteria and conducted visual inspection of time-height profiles for every DCC core included in the analysis to ensure data from the storm edge are not included.

We have added the following text in Section 2.3: "In Figure 1a and 1b, we highlight the cell tracks for isolated DCCs from the wet (December to April) and dry season (June to September). In Fig. 2, we show wind rose diagrams based on the propagation direction of these cells and heatmaps from a 2D histogram of latitude and longitude pairings from the cell tracks."

We have added the following text in Section 3.1.2: "There were minor differences across seasons in terms of the distance of the tracked cell from the MAO site. About 39% and 33% of the tracks were within 0.1-degree latitude and longitude of the MAO site and about 80% and 77% of the tracks were within 0.2-degree latitude and longitude of the MAO site for the dry and wet season, respectively (Fig. 2b, d)."

Tracking storms with a coarse time-step. The authors focus on tracking isolated cells with a 12-minute radar data frequency at 2 km AGL. Could there be instances in the dataset where a storms dissipates, and a new one forms in similar location within the 12-minute time step, which is therefore tracked as a continuous storm? This seems like a large source of uncertainty that isn't discussed.

We agree that the coarse temporal resolution of the SIPAM radar can lead to uncertainties in the cell tracking output. However, our use of the SIPAM radar data for cell tracking is dictated by the lack of observations with higher temporal resolution during the GoAmazon2014/5 field campaign. Multiple other studies have also used the SIPAM radar data with the 12-minute resolution to conduct cell tracking of isolated deep convection (Giangrande et al., 2023) and examined the uncertainties associated with different thresholds applied to cell tracking (Leal et al., 2022).

We have added the following text to the first paragraph in Section 2.2 to acknowledge this limitation: "It is important to note that the cell tracking outputs heavily depend on the temporal resolution of the input data. Unfortunately, the SIPAM radar, being part of the Amazonian operational weather radar network, uses a fixed temporal resolution of 12 minutes that could not be changed (Saraiva et al., 2016). Ideally, a smaller repetition time would result in a more accurate description of the convective processes. Nevertheless, this 12-minute temporal resolution is better than the satellite alternative (30 minutes)."

To objectively test the hypothesis of cell replacement and to quantify this uncertainty, we need to compare cell tracking output using SIPAM or GOES-13 with cell tracking output from using a faster scanning radar (not available for the region and time period of this study) or a newer satellite like GOES-16. However, this is outside the scope of this current work. Our goal is to conduct a follow-on study of aerosol-cloud interactions in the region. For this objective, we need to use aerosol observations and radiosonde launches conducted at the MAO site during the GoAmazon2014/5 field campaign. While it would certainly help to use data with better temporal resolution from newer datasets, such observations would not be temporally co-located with measurements from GoAmazon2014/5. Our group is conducting cell tracking in locations like Houston, Texas, where observations with better spatial and temporal resolution are available. A thorough comparison of cell tracking outputs with data input with variable temporal resolution is left for future work.

Splits/mergers. What happens if a storm splits into two storms or merges with a different storm? Could these events be present in the dataset? This is related to concerns about the coarse time step for tracking, and how confident the authors can be about the determination of lifecycle stages.

Our dataset examines "ISO" days identified by Giangrande et al. (2023). Their methodology for selecting the "ISO" days required the SIPAM radar to sample a DCC for over 48 mins in the radar domain without an obvious split/merger. In addition, isolated deep convective cells tend to merge/split much less relative to larger, mesoscale convective systems which are removed from this analysis. A recent study used the SIPAM radar data from the first year of GoAmazon2014/5 and tested the impact of various thresholds on cell tracking outputs (Leal et al., 2022). The authors found that continuous features (features without splits/mergers) represent most of the tracked cases and a feature identification threshold of 30 dBZ minimizes the proportion of features with splits/mergers. Based on these factors and our minimum feature identification threshold of 30 dBZ, splits/mergers should have minimal influence on our study. A splits/merger functionality was recently added to the *tobac* algorithm and we plan to test the function in our future work.

We have added the following text in Section 2.2: "The minimum identification threshold of 30 dBZ was found to minimize the proportion of tracked systems with mergers or splits in the Amazon to under 20 % (Leal et al., 2022)."

Moreover, the utilization of statistical trends in the convection properties helps us confidently attribute the lifecycle stages to the cell tracking outputs. This confidence is further reflected in the consistency across observations from independent instruments, i.e., observations of the highest rainfall rates from a disdrometer and strongest updrafts from the RWP for cells defined as "mature" based on SIPAM radar data and *tobac* cell tracking.

Using 2 km reflectivity. The authors use 2 km reflectivity for tracking and defining feature centers (L127). Why do they use this altitude, as opposed to a column maximum reflectivity that may be more accurately located with the updraft? Have they tried using different altitudes? For storms that may not be precipitating heavily, could the 2 km reflectivity with a 30 dBZ minimum threshold miss some of the beginning and ending stages of a storm lifetime? Similarly, for sheared storms, the 2 km location may not accurately describe the updraft location?

Locating updrafts while tracking reflectivity features: In this study, the goal of using *tobac* was to identify the lifecycle stage of convection sampled by the RWP at the MAO site. The precise location of the updraft relative to the maximum reflectivity within the tracked features should not impact the estimation of the lifecycle stages. Following the maximum reflectivity within the feature allows us to determine the age of a feature when it is closest to the MAO location. Since this is the extent of the contribution of *tobac* cell tracking to the analysis of updraft or downdraft cores, our methodology is sufficient as it relates to using reflectivity from a single altitude rather than the column maximum reflectivity. Similarly, from the perspective of sheared storms, the tilt in an updraft should not impact the estimated age of the cell when it is sampled by the RWP. Furthermore, previous studies have shown the 0-5 km wind shear to be lower at MAO compared to other convective environments like the southern great plains in the USA (e.g., Wang et al., 2019). Our selection criteria in Section 2.5 rejects instances when cell tracking identified an isolated cell near the RWP but the updraft was incompletely sampled by the RWP. This should account for any uncertainties in locating updrafts using the cell tracking algorithm.

We have added the following text to Section 3.3: "It is assumed the RWP sufficiently sampled the updraft core of the DCC at this timestep as the RWP sampling for the DCC passed our selection criteria in Section 2.5."

Cell tracking using reflectivity at different altitudes: We examined CAPPIs for an altitude of 1 km but the reflectivity field at this altitude was prone to the inclusion of ground clutter and radar artifacts. We used a higher altitude of 2.5 km as a sensitivity test for certain cases and this did not have a substantial impact on the output. We did not track cells using altitudes of 3 km or higher due to the potential for the exclusion of developing or dissipating stages of convection at these altitudes.

Using a 30 dBZ minimum threshold: This threshold was chosen based on the visual inspection of the radar data. The radar data was prone to artifacts with reflectivity values up to 20 dBZ. To avoid the tracking of such artifacts, we chose a threshold of 30 dBZ. Giangrande et al. (2023) tracked cells with thresholds of 25 dBZ and 35 dBZ and found similar bulk statistical trends as it relates to the cell tracking and when comparing tracked convection from different seasons. Finally, as mentioned earlier, the selection of this threshold limits the impact of splits/mergers for this study region (Leal et al., 2022).

To address the previous two comments, we had added the following text to Section 2.2: "A minimum threshold of 30 dBZ and the altitude of 2 km for the CAPPI data were chosen following the visual inspection of the radar data which revealed ground clutter and artifacts at lower altitudes and with reflectivity values up to 20 dBZ. The selection of 30 dBZ as the minimum threshold also ensures the minimization of splits/mergers in our dataset (Leal et al., 2022)."

Using reflectivity at 2 km for cell tracking: The choice of selecting a single altitude to represent the updrafts was dictated by the limitation of the cell tracking algorithm at the time of conducting data analysis. *tobac* has been updated since then to include three-dimensional cell tracking, but the functionality was added fairly close to the time of submission of this manuscript. The *tobac* manuscript that describes the latest tobac version is currently under peer review at the time of writing. Lines 722-725 of the original manuscript acknowledged that the open-source tracking algorithms are being updated and that these updates will allow us to conduct three-dimensional cell tracking in future work. Three-dimensional cell tracking is unlikely to affect the age of the convection when sampled by the RWP at MAO (as discussed above). The dependence of our results on the variability in cell area as a function of height is minimized by the use of p(w) while calculating mass flux and transport rate (further discussion below).

p(w) seems to play an important role in the calculation of vertical mass flux and transport but is placed in the supplement without much discussion of how it is calculated and what it means. The authors do include a brief explanation of p(w), but it was not clear (L460-462). Can the authors spend more time describing this, given its importance in their calculations of mass flux?

- L448-450: The authors state that 2 min periods of RWP observations are selected, and this is based on an assumed median updraft width of 1 km and average propagation speed. I do not understand how these variables are related to the 2 min period of RWP observations that are chosen, and found this sentence confusing? Can the authors make this clearer?
- L452: The authors state that the average updraft/downdraft are weighted by the probability of sampling an updraft/downdraft during this time period (p(w)). However, the authors do not provide a clear description of how p(w) is calculated and place the figure of p(w) in the appendix (Figure S3). Can the authors provide a clearer description of how p(w) is calculated, and place this figure in the main manuscript, given its importance in calculating mass flux?

We have moved Figure S3 from the original supplementary materials to the revised manuscript as Figure 13. We have added a new figure (Figure 12) that shows composite time-height profiles from the RWP. Specifically, the composite vertical velocity and reflectivity are shown, where the composite values represent median values over all DCCs classified as “mature”. Composite data for the other four lifecycle stages are provided as Figure S1 in the revised supplement.

Given previous observations of a median updraft width of 1 km for MAO (Wang et al., 2020), we use a 2-min period of RWP observations centered at time 0 to represent the vertical air motion within each DCC core (Figure 12 shows data up to 10 mins before or after time 0). The use of an average value for updraft or downdraft speed over the 2-min period also reduces the influence of a single profile of RWP measurements from time 0. Natural variability in the shape of updraft or downdraft regions within the DCC core can lead to variations in the vertical velocity profiles over the 2-min period. To account for these natural variations, the time series of vertical velocity over the 2-min period is weighted by the probability, p(w), of sampling an updraft or a downdraft during the 2-min period. For each height level, the value of p(w) for updrafts (or downdrafts) is determined as the ratio of the number of observations with $w > 1$ m s^{-1} (or $w < -1$ m s^{-1}) and the total number of observations over the 2-min period (20 observations given the RWP resolution of 6 seconds).

We have added the following text to answer the referee’s questions and to avoid confusion regarding the calculation and utility of p(w):

“The RWP data are used to characterize the properties of the DCC cores. Time-height profiles of the composite w and Z for DCCs classified as mature DCCs are shown in Fig. 12 with composites for other lifecycle stages in Fig. S3. These composites represent the median values of w and Z across all DCCs classified within the lifecycle stage. The profiles are centered at time 0 which represents the time of simultaneous sampling by the SIPAM radar and the RWP during the DCC overpass at MAO. This is the timestep for which we have the DCC lifecycle classification from the cell tracking. Given an average propagation speed of 9 m s^{-1} (Fig. 4), a 2-min period of RWP observations corresponds to sampling a core that is 1.08 km wide. Based on previous observations of a median updraft width of 1 km for MAO (Wang et al., 2020), we use a 2-min period of RWP observations centered at time 0 to represent each DCC core. The use of an average value for updraft or downdraft speed over the 2-min period also reduces the influence of a single profile of RWP measurements from time 0.

Natural variability in the shape of updraft or downdraft regions within the DCC core can lead to variability in the vertical velocity profiles over the 2-min period. To account for these natural variations, the time series of vertical velocity used to represent the DCC core is weighted by the probability, p(w), of sampling an updraft or a downdraft during the 2-min period. For each height level, the value of p(w) for updrafts (or downdrafts) is determined as the ratio of the number of observations with $w > 1$ m s^{-1} (or $w < -1$ m s^{-1}) and the total number of observations over the 2-min period (20 observations given the RWP resolution of 6 seconds).”

Additional Methodology Clarifications:

L164: How exactly are isolated convection (“ISO”) days determined? The authors point to Giangrande et al., 2023, but should provide the description here, given that this is defining the dataset used in this study.

We have added the following text here: “Their methodology required the SIPAM radar to sample a DCC for over 48 mins in the radar domain without an obvious split/merger and without systems larger than 1000 km^2 which could represent mesoscale convective systems or squall lines.”

The authors state that isolated DCCs observed on “ISO” days are used for subsequent analysis. How are these isolated DCCs chosen? I am especially curious given that some 500 km2 storms seem to make it into their dataset (i.e., L274).

After identifying the “ISO” days, as a first step, we examined the cells within 100 km of the SIPAM radar which also reached within 20 km of the MAO site. Giangrande et al. (2023) applied an area threshold of 1000 km^2 while selecting the ISO days. This meant cells with area greater than 500 km^2 could make it into the initial dataset. However, as pointed out earlier, not all of the cells identified on the ISO days were sufficiently sampled by the RWP at MAO. Such cells were removed from subsequent analysis if the RWP data from the corresponding time period did not indicate sufficient sampling of the updraft core (see selection criteria in Section 2.5).

In certain instances, cells with areas between 500 and 1000 km^2 could indicate large but isolated convection. Previous studies have used much higher thresholds to define organized convection.

For example, Paccini and Stevens (2023) used a threshold of 10,000 km^2 to represent organized convection over the Amazon. In a review of tropical convection, Houze et al. (2015) defined echo objects with an intensity of 30 dBZ to have organized into mesoscale systems if the objects had area greater than 800 km^2.

It seems like many of the tracked storms last 3 or 4 time steps, but how can the authors classify these cells into their 5 lifecycle bins?

We addressed this comment on pages 2-3 of this document in response to a comment by referee #1.

How big is the radar domain and is this constrained by using 2 km CAPPIs, given that the radar beams will increase in altitude away from the radar?

Saraiva et al. (2016) describe the operations of the Sistema de Proteção da Amazônia (SIPAM) weather radar network. The SIPAM radar has a beamwidth of 1.8° and performs two volumetric radar scans every 12 minutes. The first scan covers a domain of 240 km from the radar location with gate resolution of 500 m, azimuth resolution of 1°, and 17 elevation angles (0.9° to 19.5°). The second scan covers 400 km with three elevation angles (0.9° to 3.7°). Given these parameters, Saraiva et al. (2016) limited the 3D representation of the radar data and computed CAPPIs up to a distance of 150 km from the radar location. In comparison, this study uses a stricter threshold of 100 km from the radar location. Our data analysis further requires the cells to be sampled within 20 km of the MAO site location with sufficient sampling of the updraft core by the RWP located at the MAO site.

Previous studies have used gridded SIPAM radar data to evaluate satellite rainfall retrievals and model simulations (e.g., Oliveira et al., 2016; Barber et al., 2022). Oliveira et al. (2016) compared rainfall retrievals derived using SIPAM radar data over 100 km away from the radar location with an X-band radar and found the two radars showed similar results. These comparisons combined with our use of a stricter threshold of 100 km for data selection gives us confidence in using the radar data.

We have added the following text to Section 2.1: "The SIPAM radar has a beamwidth of 1.8° and performs two volumetric radar scans every 12 minutes. The first scan covers a domain of 240 km from the radar location with a gate resolution of 500 m, azimuth resolution of 1°, and 17 elevation angles (0.9° to 19.5°). The second scan covers 400 km with three elevation angles (0.9° to 3.7°). Given these radar configuration parameters, Saraiva et al. (2016) limited the 3D representation of the radar data only up to 150 km from the radar location and computed Constant Altitude Plan Position Indicators (CAPPI) for the 150 km domain. In comparison, this study uses a stricter threshold of 100 km from the radar location. The clutter-corrected SIPAM Z was gridded onto a 1 x 1 km grid for 2 km CAPPIs."

Can storms that initiate outside the radar domain move into the domain? If so, how would the authors know its lifecycle stage?

As shown in Figure 1 in the revised manuscript, we do not consider cells that are more than 100 km from the radar and not sampled within 20 km of the MAO site during their lifetime. Given that we are tracking isolated convection, it is unlikely a cell forms outside the radar domain (400 km) and reaches within 20 km of MAO.

"Average Z:" Many times in the manuscript the authors talk about average Z, and sometimes I was confused whether they are referring to the 2 km CAPPI Z used in the tracking, Z from the SIPAM radar at different altitudes, or the Z from the RWP observations. Can the authors make this clearer throughout the manuscript?

We did not use Z from the SIPAM radar at different altitudes in this study. Throughout the manuscript, we now refer to Z from the SIPAM radar data as "SIPAM Z" or "Z from the SIPAM radar". References to Z from the RWP are left as "Z" as they were in the original manuscript.

L234: Does "at least 10 consecutive cloud echoes" mean consecutive in time or height?

We meant consecutive in height. We have modified the text to avoid this confusion.

L366: When the authors use the term "Mature" DCCs (here and elsewhere in the manuscript), are the authors referring to data in lifetime bin 3 or data in lifetime bins 2-4, since bins 2 and 4 are defined as "early mature" and "late mature?"

Mature DCCs only refer to data from lifetime bin 3. We have modified the text in Section 2.4 to the following to avoid confusion: "The 3rd lifetime bin, when DCCs reach their peak Z and A, is defined as the mature stage (Figs. 3, 5, 6). Following this definition, data for "mature DCCs" corresponds to data from the 3rd lifetime bin throughout the study."

L556: The authors state that "time 0 represents the timestep when the indicated lifecycle stage of the DCC was identified using tobac." I am having a hard time interpreting what this means. What if there are more than one radar timesteps for a given lifecycle stage? In that case, which time is used?

We have updated this sentence as follows: "time 0 represents the timestep when the DCC was simultaneously sampled by the SIPAM radar and the RWP at the MAO site. This is the timestep for which we have the DCC lifecycle stage determined from the cell tracking. It is assumed the RWP sufficiently sampled the updraft core of the DCC at this timestep as the RWP sampling for the DCC passed our selection criteria in Section 2.5."

For each timestep, only one reflectivity feature was used to assign the lifecycle stage to the RWP data. This feature was chosen based on its distance from the MAO site. If there were more than one radar timesteps within a given lifecycle stage, the timestep with the feature closest to the MAO site was chosen.

L556: Is the equivalent potential temperature from MAO used even when cells are >20km away from MAO? If so, can what is happening at MAO be accurately related to the storm dynamics from the respective storm?

The feature position from cell tracking is not used to represent the updraft properties. The updraft properties at MAO are derived solely from RWP data. In Section 2.5, we conduct extensive checks and apply strict thresholds to make sure the RWP at the MAO site sampled a strong ($w > 1$ m s^{-1}) and continuous updraft core within an isolated convection event. We are confident that the equivalent potential temperature represents the storm dynamics as they were sampled by the RWP.

L609: What Tb measurement is being used in this analysis? Is it the Tb measurement at MAO, at the feature center, or some average over the entire feature?

Considering your comment below, the length of the manuscript, and the addition of text and figures in response to referee comments, we have removed section 3.4 from the manuscript.

Figure 1 (L762): Is this data for all tracked cells, or only the cells that go into the analyses presented in the remaining figures?

This is data for cells that go into the analyses presented in the remaining figures.

Figure 2 (L768): Are "other DCC days" the same as ACE days as described in the text? If so, can the authors use the same terminology?

We have updated the legend in Fig. 2 to remain consistent with the text.

Figure 2: (L770): Why is there a propagation speed threshold here of 0.5 m/s? I don't believe this was mentioned in the text?

We have added the following text to Section 2.3: "As a final check, we remove any cells with propagation speed below 0.5 m s^{-1} to avoid the inclusion of ground clutter or radar artifacts."

Figure 3: For this (and all) boxplot figures, what is being shown? For example, do the bars that extend out represent the max/min values or some other value? A clearer description of that is being shown in the boxplots is necessary.

We have added the following text to the caption for Figure 3 for clarity: "The box lengths represent the interquartile range and the whiskers extend to the 5th and 95th percentiles."

Figure 16: Is this showing the mean or median equivalent potential temperature?

This is the mean equivalent potential temperature.

L826: Is there a typo here? Should Figure 8 be Figure 13?

The figure caption has been updated.

There is an inconsistency of Figure S3 caption to what is stated in the figure image (within 1 min versus within 2 mins).

We have placed the image in the revised manuscript and updated this typo in the caption.

Other Suggestions:

L69: The NASA Investigation of Convective Updrafts Mission (e.g., van den Heever, 2021; Stephens et al., 2020 Prasanth et al., 2023) is planning to do this, so it would probably be good to mention this here, since this research has implications for this mission as well.

We have added the following text in Section 1: "Ground-based observations of the air motions within convective clouds can complement satellite missions aiming to quantify the vertical mass transport in convective storms, for example, the NASA Investigation of Convective Updrafts (INCUS) mission (Stephens et al., 2020; van den Heever, 2021; Prasanth et al., 2023)."

Many of the results involve comparing results between the different seasons, which show up in different figures and makes comparing them quite difficult. Can the authors combine figures effectively, so that differences between results can be more clearly assessed? This is particularly the case for Figures 12-14 and 8-10.

We have updated the figures to plot data from dry and wet seasons together in one figure.

In terms of assessing whether their smaller sample had sampling biases, what is the reason for using the nearest neighbor distance as a metric, as opposed to metrics (i.e., reflectivities, ETHs) that may be more telling of the types/intensities of storms being sampled? It seems like it would be important to understand if their sample represents both weak and strong storms, compared to the possible population of storms?

The nearest neighbor distances were used to compare the two datasets to highlight the degree of isolation of convection on ISO and ACE days and the limited domain of the data used here compared to the entire radar domain. Since the focus of our study is on isolated convection, we think this is an important result to confirm the hypothesis that the cells examined in our study are unlikely to be affected by surrounding convection and if there is any influence, it is representative of the region.

L457: For calculating mass flux, the authors assume the storm area is constant as a function of height, based on the area of the storm at 2 km AGL? It seems like the authors potentially have the data to validate this assumption (i.e., SIPAM radar data), or can find supporting literature for this assumption?

The impact of this assumption is limited using p(w) while calculating mass flux and mass transport rate. The width of updrafts/downdrafts would vary with height, and this is captured by the probability of sampling updrafts/downdrafts during the sampling duration from the RWP. With the availability of 3D cell tracking from *tobac* 1.5, we leave the examination of the direct impact of this calculation to future work when 3D cell tracking will be conducted.

This manuscript is already quite lengthy, and while all the analyses are interesting, perhaps some areas (i.e., S3.3) are less relevant to the authors focus on updrafts and can be removed. This may allow them to better describe their other analyses within a more reasonable space.

After considering the referee's comment and the length of the revised manuscript, we have removed section 3.4.

References:

van den Heever, S. (2021). NASA selects new mission to study storms, impacts on climate models. NASA Earth. Retrieved from https://www.nasa.gov/press-release/nasa-selects-new-mission-to-study-storms-impacts-on-climate-models

Stephens, G., van den Heever, S. C., Haddad, Z. S., Posselt, D. J., Storer, R. L., Grant, L. D., et al. (2020). A distributed small satellite approach for measuring convective transports in the Earth's atmosphere. IEEE Transactions on Geoscience and Remote Sensing, 58(1), 4–13. https://doi.org/10.1109/TGRS.2019.2918090

Prasanth, S., Haddad, Z. S., Sawaya, R. C., Sy, O. O., van den Heever, M., Narayana Rao, T., & Hristova-Veleva, S. (2023). Quantifying the vertical transport in convective storms using time sequences of radar reflectivity observations. Journal of Geophysical Research: Atmospheres, 128, e2022JD037701. https://doi.org/10.1029/2022JD037701.

Barber, K. A., Burleyson, C. D., Feng, Z., & Hagos, S. M.: The Influence of Shallow Cloud Populations on Transitions to Deep Convection in the Amazon. J. Atmos. Sci., 79(3), 723-743, 2022.

Oliveira, R., V. Maggioni, D. Vila, and C. Morales, 2016: Characteristics and diurnal cycle of GPM rainfall estimates over the Central Amazon region. Remote Sens. 8, 544, https://doi.org/10.3390/rs8070544.

Saraiva, I., Silva Dias, M. A. F., Morales, C. A. R., and Saraiva, J. M. B.: Regional variability of rainclouds in the Amazon basin seen by a network of weather radars, J. Appl. Meteorol. Clim., 55, 2657–2675, https://doi.org/10.1175/JAMC-D-15-0183.1, 2016.

Houze, R. A., Jr., K. L. Rasmussen, M. D. Zuluaga, and S. R. Brodzik (2015), The variable nature of convection in the tropics and subtropics: A legacy of 16 years of the Tropical Rainfall Measuring Mission satellite, Rev. Geophys. , 53, 994–1021, doi:10.1002/2015RG000488.

Paccini, L., & Stevens, B. (2023). Assessing precipitation over the Amazon basin as simulated by a storm-resolving model. Journal of Geophysical Research: Atmospheres, 128, e2022JD037436. https://doi.org/10.1029/2022JD037436.